ELSEVIER
CrossMark
Available online at www.sciencedirect.com
ScienceDirect
Procedia Computer Science 90 (2016) 132 – 137
Procedia Computer Science

International Conference On Medical Imaging Understanding and Analysis 2016, MIUA 2016, 6-8 July 2016, Loughborough, UK

Dots detection in HER2 FISH images based on alternative color spaces

Tomasz Les[a,*], Tomasz Markiewicz[a,b], Marzena Jesiotr[b], Wojciech Kozlowski[b]

[a]*Warsaw University of Technology*, [b]*Military Institute of Medicine, Warsaw, Poland*

Abstract

This article presents an algorithm for detection of HER2 gene and chromosome centromere - CEN17 signals in FISH images. FISH is a popular fluorescence staining method which allows the evaluation the degree of HER2 gene amplification on breast cancer tissues. The FISH test allows the choice of an appropriate treatment. To ensure the best discrimination of the spots from the non-spots elements (i.e. a stroma) the algorithm was based on the color spaces transformation. The main point of this algorithm is an automatic red and green spots detection using transformations into different color spaces. Quality-identification tests were performed by using one-class SVM. The selection of an optimal color space for identification of the spots was completed using automatic Fisher feature selection. The results of the experiments are promising and show good effect of color transformation application on genes recognition system. Detailed results are presented.

Peer-review under responsibility of the Organizing Committee of MIUA 2016

Keywords: Image processing; color space; one-class SVM; FISH

1. Introduction

Breast cancer is the most common type of invasive cancer amongst women. Its mortality is the first or second (closely after lung cancer) from cancers in the most countries. Its treatment may include surgery, medications (hormonal therapy and chemotherapy), radiotherapy and/or immunotherapy. Histological diagnosis of breast cancer is necessary in the surgical intervention, cancer classification, prognosis and, most of all, treatment selection. One such evaluation is a visualization of HER2 gene expression, based on fluorescence in situ hybridization (FISH)

* Corresponding author.
E-mail address: lest@ee.pw.edu.pl

Peer-review under responsibility of the Organizing Committee of MIUA 2016
doi:10.1016/j.procs.2016.07.025

cytogenetic technique. This method is critical for the trastuzumab therapy application, when the immunohistochemical staining does not give the explicit indication. The FISH technique allows us to visualize the HER2 gene and chromosome centromere (CEN17) in the cancer cells in order to detect the HER2 gene amplification. Only in that case the trastuzumab therapy can be successful and should be applied.

The fluorescence technique allows us to visualize selected genes in the fluorescence microscope. In the breast cancer diagnosis, the DAPI filter for nuclei, FITC for centromeres and Texas Red for HER2 genes are used. After applying these filters, centromers become green, HER2 genes become red and nuclei become blue. The eye colors are in accordance with the RGB color space, which allows to create a color image and, moreover, RGB colorspace is well recognizable by the human eye. This results in the fact that most computer methods presented in the literature are based on that color representation. The blue channel is used for the nuclei recognition, while red and green channels are used independently for spots recognition [2,3,4]. The different compositions of green and red channels can be used for the better dot extraction from the image [4]. However in pathology practice, the differentiation of the color shades in specimens is still a real, unsolved problem. The question of nuclei detection is sometimes avoided5, but the dot detection is the most important step in the FISH image analysis.

In this paper we examined various color spaces for dots detection, with the aim to find the best channel representation. Together with them, we applied one-class Support Vector Machine (SVM) for delimitation of the true dots colors. This is the most important for the centromere dots which have a similar color to the stroma, sometimes located closely to nuclei region or even stroma overlaps nuclei which gives false green signals.

2. Material and methods

We based our study on twelve patient cases from the archive of the Department of Pathomorphology, Military Institute of Medicine, Warsaw, Poland. All of them are classified as 2+, basedon immunihistochemical membrane HER2 staining and are referred to further diagnosis based on FISH. The specimens were stained by Dako HER2/neu test and the images were acquired on an Olympus BX-61 microscope with the fluorescence selective filters and DP-72 camera, cellF software, at the magnification 1000x and resolution 2070x1548. Blue, green and red channels were used as the components to make the analyzed RGB images.

In Fig. 1 FISH images are presented used in the diagnosis of breast cancer. The main task is to detect and count the red and green spots in the images.

a b

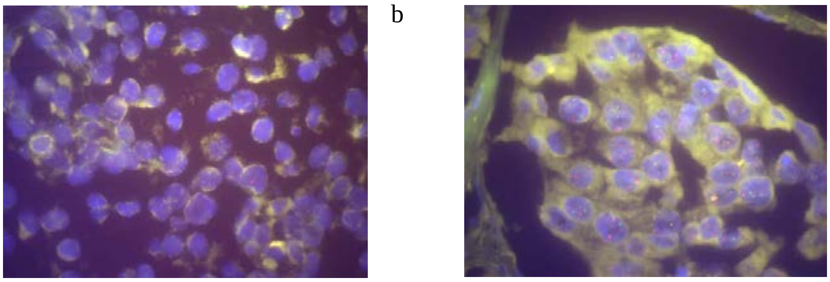

Fig. 1 (a) The example of a FISH image without gene HER 2 over expression;
(b) The example of a FISH image with gene HER 2 over-expression.

2.1. Color spaces

The RBG color space is commonly used for storing the FISH images. This color space is adequate for manual evaluation, but it limits the possibility of an automatic recognition of some image components. Especially, the green-yellow-like stroma is similar to the centromere dots which make the automatic spots detection more difficult. Often, the HER2 red dots are also barely distinguishable in the nucleus area. Color thresholding allows us to recognize dots but in fact points are often counted redundantly with this method. Thus, we considered the possibility of using the following color spaces: Hue-Saturation-Value (HSV), Hue-Saturation-Lightness (HSL), CIE L*u*v*

(CIELUV), CIE L*C*H* (CIELCH) and YCbCr - "digitized" version of YPbPr. Each model provides new, different values for image analysis. Five transformations of the RGB model into other alternative models have been defined. An example of a dot with possible transformations is shown in Fig 2.

RGB -> HSV
RGB -> HSL
RGB -> CIELUV
RGB -> CIELCH
RGB -> YcBcR

Fig. 2 Example of green spot and possible transformations to alternative color spaces.

The same image has been converted into other spaces: HSV, HSL, CIELUV, CIELCH and YCbCr. In Fig 3 there are shown characteristics of the selected channels luminosity function after transformation into other color spaces for part of the view from the previous image.

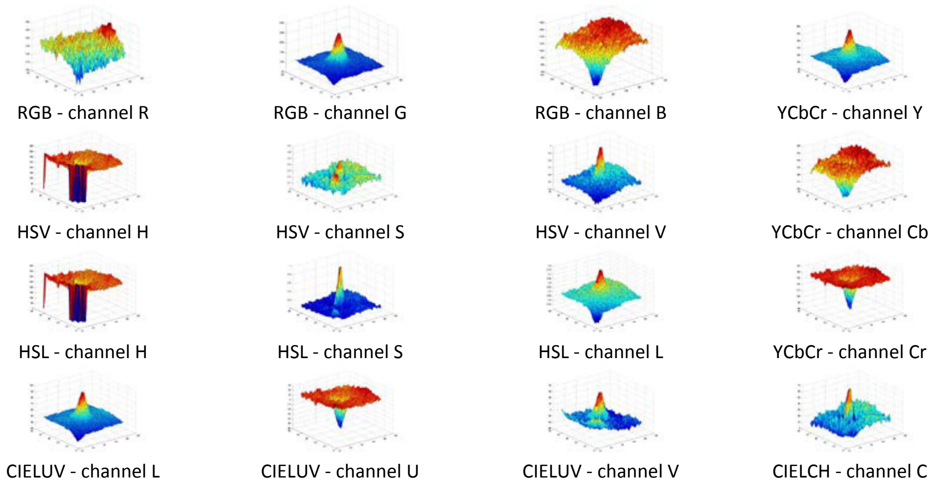

Fig. 3 The values of the Fisher discriminant measure for the selected 16 channels in six colors spaces for green spots.

The transformation brings new important information, which can simplify the development of image recognition methods and increase their efficiency.

2.2. Color channel discrimination

The color channels generated from the six color spaces should be evaluated in respect of their utility to the dot recognition. For this task, the Fisher discriminant metric was applied[6]. The measure of the quality of the features for a pair of classes can be expressed in the form:

$$S_{ij}(f) = \frac{|\mu(f_i) - \mu(f_j)|}{\sigma(f_i) + \sigma(f_j)} \tag{1}$$

where μ is a mean and σ is a standard deviation. The *S* value calculated for each feature informs which of them differentiate best the selected classes.

2.3. One-class Support Vector Machine as classifier

The determination of a set of most utility color channels provides the n-dimensional feature space, that can be the input for the classifier. The role of the classifier is to find the most representative region in space for pixels belonging to the specific dots: one region for the CEN17 dots, second for the HER2 dots. The aim is to recognize only those regions. Moreover, the representative data are available only for those areas, so the application of the one-class SVM is reasoned. One-class SVM is a specific type of SVM, whose aim is to separate all initial data points from the origin (feature space) and to maximize the distance from the border hyperplane to the origin. Application of the SVM classifier allows us to find all pixels forming the markers (red or green). Grouping the pixels into clusters, which form the final recognized spots requires an application of k-nearest neighbour technique. If all pixels of a cluster are located inside the circle of the predefined radius (in experiments 20 pixels radius), then a cluster is classified as a spot.

3. Results

In numerical results we considered 18 potential descriptors (each color space contains three channels) from six colors spaces, as shown in table 1. Figure 4 presents Fisher measure values (obtained on the basis of the formula 1) in accordance with the numbers of channels.

Table 1. Numbers of the channels from different color spaces.

Channel identifiers for each colour space	Color space
1| 2 | 3	R | G | B
4 | 5 | 6	H | S | V
7 | 8 | 9	H | S | L
10 | 11 | 12	L | U | V
13 | 14 | 15	L | C | H
16 | 17 | 18	Y | Cb | Cr

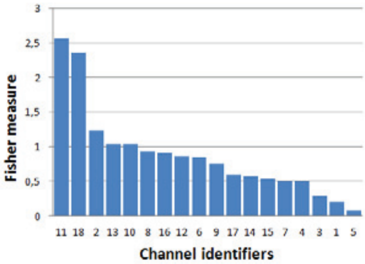

Figure 4. Values of the Fisher discriminant measure for the 18 channels in six colors spaces for red spots.

As is seen, the highest cumulative value represents the features number: 11 (LUV - channel U) and 18 (YcbCr - channel Cr). Figure 5 presents pixels distribution of red spots (the red points on the graph) and other non-red areas (the black points on the graph) marked by a human. The following distribution shows the best and the worst red/non-red pixels separation.

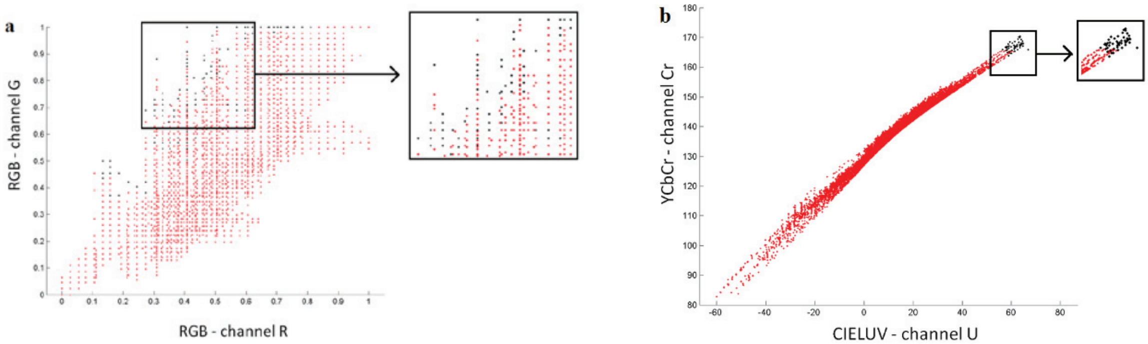

Fig. 5 Distribution of red spots (the red points) and non-red elements (the black points). A- channel R and G of RGB color space. B- channel U and Cr after transformation. In squares are visible enlarged fragments of points distribution.

Similar results have been obtained by analysing green spots using Fisher measure. In the experiments we used 12 FISH image cases - 6 with HER2 gene over-expression and 6 images without gene HER 2 over expression. The quality of red and green spots detection system (AS) was examined in comparison to the detection of a human expert (Hum). As training data sets for SVM we used pixels which form 35 red and 35 green spots after LUV channel U and YCbCr channel Cr transformation. Table 2 presents the results of the automatic spots detection.

Table 2. Test results of red and green spots detection

Case:	Hum red	Hum green	AS red	AS green	Error red	Error green
1	57	63	48	60	**15.79**	**4.76**
2	59	42	49	40	**16.95**	**4.76**
3	89	92	79	88	**11.24**	**4.35**
4	73	74	62	67	**15.07**	**9.46**
5	117	50	103	44	**11.97**	**12.00**
6	136	59	122	53	**10.29**	**10.17**
7	51	49	46	47	**9.80**	**4.08**
8	35	36	30	33	**14.29**	**8.33**
9	184	79	160	70	**13.04**	**11.39**
10	121	86	100	75	**17.36**	**12.79**
11	152	56	135	50	**11.18**	**10.71**
12	125	54	111	47	**11.20**	**12.96**
				Avg. error:	**13.18**	**8.81**

Final error of red and green spots detection is defined as

$$Error = 100 - \left(\frac{AS}{Hum}\right) * 100 \quad (2)$$

The results of the automatic red and green spots detection algorithm are shown in Fig. 6.

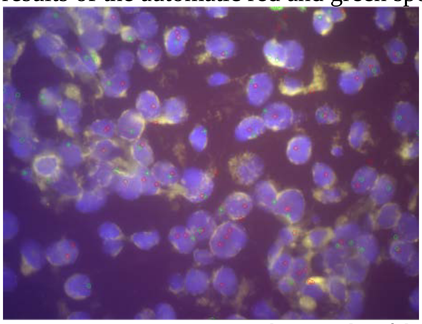
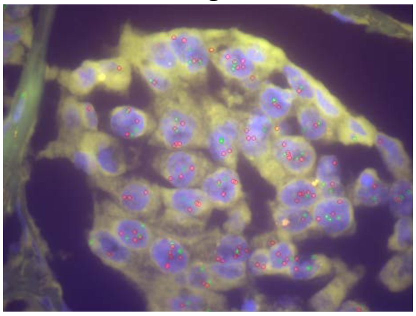

Fig. 6 The examples of the detected red and green spots in the FISH images.

As is seen in figure 6, most of the spots of both types (red and green) were recognized correctly by the automatic system. Stroma, whose color is similar to the green spots, is correctly not classified as a marker. Good features of matching the SVM to learning data in combination with the color space transformation allows differentiation elements with similar but different color.

4. Discussion

Counting the signals in breast cancer cells is a main stage in HER2 gene amplification evaluation. In the last two decades, a lot of approaches were developed to the problem of the FISH spots detection in the images[1-4,7-10]. The morphological image analysis, such as top-hat transform[7,8], watershed method[9], were applied. Also, cross-correlation and gradient evaluation were proposed[1]. Most of them are based on the original RGB color space. The HSI color space was used[10]. In our paper we extended analysis to the six color spaces, we found the most important channels for spots recognition and we applied one-class-SVM as a classifier. The obtained results with only 8.8% for green and 13.18% for red spots errors are satisfactory. Color transformations may also be used in implementation of other, custom methods for dots (or other objects) detection.

5. Conclusions

This paper presents a new approach of spots detection in FISH images. The developed algorithm uses color spaces transformations to increase recognition of red and green spots. Application of one-class SVM allows us to classify analyzed pixels as red, green or non-spot pixel and k-NN algorithm groups pixels into clusters which form final recognized spots. Application of color space transformation may be a valuable preprocessing step in other algorithms of breast cancer microscopic images analysis such e.g. map of colors[11] or application of 3-D models of spot detection[12]. However, more experiments and testing of different types of algorithms, applied on larger data sets are necessary to confirm practical and valuable effect of color space transformations.

Acknowledgment

This work has been supported by the National Science Centre (2012/07/B/ST7/01203 grant), Poland.

References

1. Raimondo F., Gavrielides M., Karayannopoulou G., Lyroudia K., Pitas I.,Kostopoulos I.: Automated evaluation of HER2/neu status in breast tissue from FISH images. IEEE Trans. on Image Process. 14; 2005. p. 1288-1298.
2. Theodosiou Z., Kasampalidis I.N., Kostropoulos I., et al.: Evaluation of FISH image analysis system on assessing HER2 amplification in breast carcinoma cases. The Breast 17; 2008 p. 82-86.
3. Theodosiou Z., Raimondo F., Garefalaki M.E., Karayannopoulou G., Lyroudia K., Pitas I.: "Fish Image Analysis System for Breast Cancer Studies" in Proc. of European Conf. on Emergent Aspects in Clinical Data Analysis (EACDA 2005), Pisa, Italy, 28-30 September; 2005
4. Markiewicz T., Cichowicz M., Kozłowski W. "Alternative colour space construction of the stain reaction in FISH image analysis for quantification of the HER2 gene amplification", Electrical Review, vol. 3a, p. 10-12, Mar. 2012
5. Lopez C., Tomas B., Korzynska A., Bosch R., Salvado M., Llobera M., Garcia-Rojo M., Alvaro T., Jaen J., Lejeume M.: Is it necessary to evaluate nuclei in HER2 FISH evaluation? American Journal of Clinical Pathology, vol. 139, p. 47-54, Jan. 2013.
6. Netten H., Young I. T., J.vanVliet L., Tanke H. J., Vroljik H, Sloos W. C.R., , FISH and Chips: Automation of fluorescent dot counting in interphase cell nuclei, Cytometry, 28(1); 1997 p. 1-10
7. Duda, R.O., Hart, P.E., Stork, P.: Pattern classification and scene analysis. Wiley, New York; 2003
8. Solorzano C. O., Santos A., Vallcorba I., Garcia-Sagredo J-M, Pozo F., Automated FISH spot counting in interphase nuclei: Statistical validation and data correction, Cytometry, 31(2); 1998 p. 93-99
9. Kozubek M., Kozubek S., Lukasova E., Mareckova A., et. Al, High-resolution cytometry of FISH dots in interphase nucleus nuclei, Cytometry, 36(4); 1999 p. 279-293
10. Lerner B., Clocksin W.F., Dhanjal S., Hulten M.A., Bishop C.M., Automatic signal classification in fluorescence in situ hybridization images, Cytometry, 43(2); 2001 p. 87-93
11. Les, T., Markiewicz, T., Osowski, S., Cichowicz, M.; Kozłowski, W. Automatic evaluation system of FISH images in breast cancer. Proc. of the 6th Int. Conf. Image Signal Process. LNCS 8509; 2014 p. 332-339
12. Les T., Markiewicz T., Osowski S., Jesiotr M.; Kozlowski W. Localization of spots in FISH images of breast cancer using 3-D shape analysis. Journal of Microscopy; 2016, doi: 10.1111/jmi.12360.